Bounded Normal Approximation in Highly Reliable Markovian Systems

Bruno Tuffin

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Bounded Normal Approximation in Highly Reliable Markovian Systems

Bruno Tuffin

N ° 3020

Octobre 1996

THÈME 1

Rapport de recherche

ISSN 0249-6399

Bounded Normal Approximation in Highly Reliable Markovian Systems

Bruno Tuffin *

Thème 1 — Réseaux et systèmes
Projet Model

Rapport de recherche n° 3020 — Octobre 1996 — 17 pages

Abstract: In this paper, we give a necessary and sufficient condition to perform a good normal approximation for the Monte Carlo evaluation of highly reliable Markovian systems. We have recourse to simulation because of the frequent huge state space in practical systems. Literature has focused on the property of bounded relative error. In the same way, we can focus on bounded normal approximation. We see that the set of systems with bounded normal approximation is (strictly) included in the set of systems with bounded relative error.

Key-words: Simulation, Normal Approximation, Markov Chains, Highly Reliable Systems.

(Résumé : tsvp)

* {Bruno.Tuffin}@irisa.fr

Unité de recherche INRIA Rennes
IRISA, Campus universitaire de Beaulieu, 35042 RENNES Cedex (France)
Téléphone : (33) 02 99 84 71 00 – Télécopie : (33) 02 99 84 71 71

Approximation Normale Bornée pour les systèmes Markoviens Hautement Fiables

Résumé : Nous donnons ici une condition nécessaire et suffisante pour obtenir une approximation normale satisfaisante pour l'évaluation, par les méthodes de Monte Carlo, des systèmes Markoviens hautement fiables. Les systèmes pratiques comportant en général un grand nombre d'états, on utilise des techniques de simulation. La littérature s'est concentrée jusqu'à présent sur la propriété d'erreur relative bornée. De la même manière, nous nous intéressons à la notion d'approximation normale bornée. Nous démontrons que l'ensemble des systèmes avec approximation normale bornée est (strictement) inclu dans l'ensemble des systèmes avec erreur relative bornée.

Mots-clé : Simulation, Approximation normale, Chaînes de Markov, Systèmes hautement fiables.

1 Introduction

Fault tolerant multi-components systems (which tolerate also fault propagation), as computer or telecommunication systems, are more and more reliable. Such systems are often represented by Markovian models. Direct computational time of dependability metrics using these models, as the MTTF (Mean Time To Failure) or the availability, is too expensive. Also, approximate numerical techniques, state lumping and unlumping [3], state aggregation and bounding [7] require considerable computer time and memory. The crude Monte Carlo simulation is inefficient because of the failure rarity. Thus we use variance reduction methods, principally importance sampling techniques. A general description of this technique can be found in [5]. In the literature, a number of schemes have been proposed for highly reliable Markovian systems. These include simple failure biasing [6], failure distance biasing [9], bias2 failure biasing [4] and failure distance biasing [1]. All these schemes increase the failure probability to reduce the variance of the estimator. Shahabuddin [9] introduced the notion of *bounded relative error*. If an estimator enjoys this property, we only need a fixed number of iterations to obtain a confidence interval having a fixed relative width no matter how rarely system failures occur. Here, we show that we have also to take care of the normal approximation. To obtain a bounded normal approximation is as important as a bounded relative error, because it justifies the validity of the confidence interval, and so the use of the method. Moreover, we prove that bounded normal approximation implies bounded relative error.

This paper is organized as follows: in section 2, we recall the model specifications described by Nakayama in [8] and the previous results. In section 3 we give the theorem for bounded normal approximation. Section 4 shows that systems with bounded normal approximation are also with bounded relative error and that the inclusion is strict. Finally we conclude in section 5.

2 Model presentation

We use the notations of [9] and [8]. A function f is said to be $o(\varepsilon^d)$ if $f(\varepsilon)/\varepsilon^d \to 0$ as $\varepsilon \to 0$. It is said to be $\Theta(\varepsilon^d)$ as $\varepsilon \to 0$ if $f(\varepsilon)/\varepsilon^d \to K$ where $K \neq 0$ is independent of ε. Similarly $f(\varepsilon) = O(\varepsilon^d)$ if $f(\varepsilon)/\varepsilon^d \to K$ as $\varepsilon \to 0$ and $\underline{O}(\varepsilon^d)$ if $f(\varepsilon) = c\varepsilon^{\bar{d}} + o(\varepsilon^{\bar{d}})$ with $c \neq 0$ and $\bar{d} \leq d$ when $\varepsilon \to 0$.

We suppose that the system has C types of components, with n_i components of type i. The total number of components is then $N = \sum_{i=1}^{C} n_i$. The system is subject to random failures and repairs with exponential laws. The model is given by a continuous time Markov chain (CTMC) $(Y_t)_{t\geq 0}$ defined on the finite state space S where each $x \in S$ gives for all $i = 1, \ldots, C$ the number of operational components (also called up components) of type i, $n_i(x)$. We label the state with all components up as $\mathbf{1}$. We suppose that this is the initial state. S is partitioned into two sub-sets U and F where U denotes the set of up states and F the set of down states. Of course, $\mathbf{1} \in U$. Failure propagations are allowed. Let $p(y; x, i)$ be the probability that, if the system is in state x and a component of type i fails, the system goes directly to state y by means of propagations. A transition (x, y) from a state x to a state y is said to be a failure transition if $\forall 1 \leq i \leq C$, $n_i(y) \leq n_i(x)$, with a strict inequality for some type i; this is denoted $y \prec x$. We define in the same way the repair transitions. Let Γ be the set of possible transitions. When we are in state x, a repair occur to state y with rate $\mu(x, y)$. A failure of a component of type i occur exponentially in state x with rate $\lambda_i(x)$. Given that failures are rare, we introduce a rarity parameter $\varepsilon > 0$, such that $\varepsilon \ll 1$ and

$$\lambda_i(x) = a_i(x)\varepsilon^{b_i(x)},$$

where $a_i(x) \geq 0$ and $b_i(x) \geq 1$ are independent of ε. In the same way we suppose that

$$p(y; x, i) = c_i(x, y)\varepsilon^{d_i(x,y)}$$

where $d_i(x, y) \geq 0$ is integer-valued, $c_i(x, y) \geq 0$ and $\sum_{y\in S} p(y; x, i) = 1$. We assume that repair rates $\mu(x, y)$ are independent of ε.

The infinitesimal generator of Y, denoted by $Q = (q(x, y))_{x,y\in S}$ is given by

$$q(x, y) = \begin{cases} \sum_{k=1}^{C} n_k(x)\lambda_k(x)p(y; x, k) & \text{if } y \prec x \\ \mu(x, y) & \text{if } y \succ x \\ 0 & \text{elsewhere} \end{cases}$$

for $x \neq y$, and $q(x, x) = \sum_{x\neq y} q(x, y)$. Let us denote by X the canonically embedded discrete time Markov chain (DTMC) and by $\boldsymbol{P}$ its transition matrix. If we call $b_0 = \min_{1\leq i\leq C} b_i(\mathbf{1})$ and if

$$b(x, y) = \begin{cases} \min_{i=1}^{C}\{b_i(x) + d_i(x, y) : n_i(x)a_i(x)p(y; x, i) > 0\} & \text{if } y \succ x \\ 0 & \text{if } y \prec x \end{cases}$$

is the exponent of the order of magnitude of the rate of the transition (x, y), we have [8] that for any $(x, y) \in \Gamma$,

$$\boldsymbol{P}(x,y) = \begin{cases} \Theta(\varepsilon^{b(x,y)}) & \text{if } x \neq \mathbf{1} \\ \Theta(\varepsilon^{b(x,y)-b_0}) & \text{if } x = \mathbf{1}. \end{cases}$$

Define Φ as the corresponding measure on the sample paths of the DTMC. It is known that the MTTF (Mean Time To Failure) can be expressed by the ratio [9]

$$MTTF = \frac{E_\Phi(\min(\tau_F, \tau_{\mathbf{1}}))}{E_\Phi(1_{[\tau_F < \tau_{\mathbf{1}}]})}, \tag{1}$$

where τ_F is the hitting time of the DTMC X to set F and $\tau_{\mathbf{1}}$ the hitting time to state $\mathbf{1}$. This performance measure is estimated by means of regenerative simulation. As the numerator is easy to estimate with naive simulation, we focus on the evaluation of $\gamma = E_\Phi(1_{[\tau_F<\tau_{\mathbf{1}}]})$ as in [8]. Importance sampling is used in [8] [9] [4]. As a matter of fact, a crude simulation is inefficient, very large sample sizes are required to achieve accurate estimators of γ as $\varepsilon \to 0$. We choose a new matrix $\boldsymbol{P}'$ and evaluate

$$\gamma = E_{\Phi'}(1_{[\tau_F<\tau_{\mathbf{1}}]}L)$$

where, for all path $(x_0, \cdots x_n)$, the likelihood function L is

$$L(x_0, \cdots x_n) = \frac{\Phi(x_0, \cdots x_n)}{\Phi'(x_0, \cdots x_n)}$$

and Φ' is the measure corresponding to matrix $\boldsymbol{P}'$. The most commonly used choices are balanced failure biasing, Bias1 failure biasing, Bias2 failure biasing and failure distance biasing.

Bias1 failure biasing associates, with any state $x \in U$, probabilities ρ_0 and $1-\rho_0$ which are allocated respectively to the failure transitions and to the repair transitions. In both sets, the new transition probabilities are proportional to their original ones. For all the importance sampling schemes, transitions from state $\mathbf{1}$ or from a state in F are not altered. Balanced failure biasing also allocates probabilities ρ_0 and $1-\rho_0$ to the individual failure and repair transitions, but now the same probability is allocated to each failure transition. Bias2 failure biasing [4] gives a higher combined probability ρ_1 to those failure corresponding to components types which have at least one of their type already failed. The probabilities are allocated proportionally to the original ones. Failure distance biasing [1] takes into account the minimum number of components which have to fail to put the system down. For each state $x \in U$, define the *failure distance* as $d(x) = \min_{y \prec x; y \in F} \left(\sum_{i=1}^{C} n_i(x) - n_i(y) \right)$ and the criticality of (x, y) as $c(x, y) = d(x) - d(y)$. The importance sampling matrix is defined as follows. A probability ρ_0 is allocated to the set of failure transitions and a probability $1 - \rho_0$ to the set of repair transitions. Then the set of failure transitions is divided into the set of transitions with criticality 0 (with probability $1 - \rho_d$) and the set with criticality strictly positive, called *set of dominant transitions* (with probability ρ_d). The subset of the latter, consisting of the transitions with the smallest criticalities, is assigned a (conditional) probability $(1 - \rho_c)$, and the subset of remaining transitions a (conditional) probability ρ_c. We repeat the last step as long as the remaining set contains transitions having different criticalities. In all those subsets, probabilities are allocated proportionally to the original ones.

We suppose that the system verifies the three following properties:

- A1: the DTMC X is irreducible on Ω.
- A2: for every state $x \neq \mathbf{1} \in \Omega$, there exists a state y such that $y \prec x$ and (x, y) is a transition of X in one step.
- A3: for each state $z \in F$, such that $(\mathbf{1}, z) \in \Gamma$, $q(\mathbf{1}, z) = o(\varepsilon^{b_0})$.

In [9], the author proves the following result.

Theorem 1 (Shahabuddin(1991)) *There exists $r \in \mathbb{N}^*$ such that*

$$\gamma = \Theta(\varepsilon^r).$$

In the same paper, the concept of bounded relative error is defined as follows:

Definition 1 ((Shahabuddin(1991))) *Define σ^2 as the variance of the estimator of γ and z_δ as the $1 - \delta/2$ quantile of the standard normal distribution. Then the relative error for a sample size I is defined by*

$$RE = z_\delta \frac{\sqrt{\sigma^2 / I}}{\gamma}.$$

We say that we have a bounded relative error if RE remains bounded as $\varepsilon \to 0$.

Let us denote by Δ_m the set of paths from $\mathbf{1}$ to F without returning to state $\mathbf{1}$ and with probability in $\Theta(\varepsilon^m)$, that is

$$\Delta_m = \{(x_0, \cdots, x_n) : \ n \geq 1, x_0 = \mathbf{1}, x_n \in F, x_i \notin \{\mathbf{1}, F\} \text{ for}$$
$$1 \leq i \leq n-1, (x_i, x_i+1) \in \Gamma \text{ and } \Phi\{(X_0, \cdots, X_{\tau_F}) = (x_0, \cdots, x_n)\} = \Theta(\varepsilon^m)\}.$$

We can then obtain a necessary and sufficient condition on the importance sampling measure to have a bounded relative error, which basically says that failures must not be excessively rare under Φ':

Theorem 2 (Nakayama(1993)) *For any importance sampling measure Φ' corresponding to a transition matrix P' , we have a bounded relative error if and only if for all $(x_0, \cdots, x_n) \in \Delta_m$, $r \leq m \leq 2r-1$,*

$$\Phi'\{(X_0, \cdots X_{\tau_F}) = (x_0, \cdots, x_n)\} = \underline{O}(\varepsilon^{2m-2r}).$$

3 Normal Approximation

Let us recall the following important result concerning the convergence speed to the normal law in the central limit theorem.

Theorem 3 (Berry-Essen) *[2] Let $\rho = E(|X - E(X)|^3)$, $\sigma^2 = E((X - E(X))^2)$, $\mathcal{N}(x)$ be the distribution of the centered and reduced normal law and F_I be the distribution of the centered and normalized sum $(X_1 + \cdots + X_I)/\sigma\sqrt{I} - E(X)$. Then, for each x and I*

$$|F_I(x) - \mathcal{N}(x)| \leq \frac{3\rho}{\sigma^3\sqrt{I}}.$$

Moreover, it is known [2] that

$$F_I(x) - \mathcal{N}(x) = \frac{\rho}{6\sigma^3\sqrt{I}}(1 - x^2)\mathcal{M}(x) + o(1/\sqrt{I}),$$

where $\mathcal{M}$ is the density of the centered and reduced normal law. The quality of the normal approximation depends then of the fraction ρ/σ^3.

Definition 2 *If ρ denote the third order moment and σ the standard deviation of the estimator of γ, we say that we have a bounded normal approximation if ρ/σ^3 is bounded when $\varepsilon \to 0$.*

Let us define now a class of importance sampling measures. This class increases the probability of each failure transition from a state $x \neq \mathbf{1}$. It is then the type of measure we are looking for.

Definition 3 *Let $\mathcal{I}$ be the class of measures Φ' corresponding to matrix $\boldsymbol{P}'$ defined as follows: for all $(\omega, y) \in \Gamma$, $\omega \neq \mathbf{1}$ and $y \prec \omega$,*

$$\text{if } \boldsymbol{P}(\omega, y) = \Theta(\varepsilon^d), \text{ then } \boldsymbol{P}'(\omega, y) = \underline{O}(\varepsilon^{d-1}).$$

Denote

$$\Delta = \bigcup_{m=r}^{\infty} \Delta_m.$$

For Φ' importance sampling measure, denote

$$\Delta_{m,k} = \{(x_0, \cdots, x_n) \in \Delta \ : \ \Phi\{(X_0, \cdots, X_{\tau_F}) = (x_0, \cdots, x_n)\} = \Theta(\varepsilon^m)$$
$$\text{and } \Phi'\{(X_0, \cdots, X_{\tau_F}) = (x_0, \cdots, x_n)\} = \Theta(\varepsilon^k)\},$$
$$\Delta'_t = \bigcup_{m,k \ m-k=t} \Delta_{m,k},$$

and

$$s = \min\left\{ j \in \mathbb{N} : \exists (x_o, \cdots, x_n) \in \Delta, \frac{\Phi^2}{\Phi'}\{(X_0, \cdots, X_{\tau_F}) = (x_0, \cdots, x_n)\} = \Theta(\varepsilon^j) \right\}.$$

Then, s is the integer such that $\sigma^2 = \Theta(\varepsilon^s)$. A necessary and sufficient condition on Φ' to obtain a bounded normal approximation is the following:

Theorem 4 *The normal approximation is bounded for a fixed number of iterations and a measure $\Phi' \in \mathcal{I}$ if and only if $\forall k, m$ such that $m - k < r$, $(x_0, \cdots, x_n) \in \Delta_{m,k}$,*

$$\Phi'\{(X_0, \cdots, X_{\tau_F}) = (x_0, \cdots, x_n)\} = \underline{O}(\varepsilon^{3m/2-3s/4})$$

(i.e. $k \leq 3m/2 - 3s/4$).

Before proving this theorem, let us demonstrate the following lemma:

Lemma 1 *If $\Phi' \in \mathcal{I}$, then*

- $|\Delta'_t| \le |S|^{tN+1} < +\infty$;
- *For* $(x_0, \cdots, x_n) \in \Delta_m \bigcap \Delta'_t (= \Delta_{m,m-t})$,

$$\Phi'\{(X_0, \cdots, X_{\tau_F}) = (x_0, \cdots, x_n)\} \le \alpha\delta^m\varepsilon^m \le \alpha\delta^t\varepsilon^t$$

  for ε small enough, where α and δ are constants independent of $(x_0, \cdots, x_n)$.
- *For* $(x_0, \cdots, x_n) \in \Delta'_t$, $L(x_0, \cdots, x_n) \le \kappa\eta^t\varepsilon^t$ *for ε sufficiently small, where η is a constant independent of ε and* $(x_0, \cdots, x_n)$.

Proof: On Δ'_t, we can not have more than t failures from a state different of $\mathbf{1}$, then not more than $t+1$ failures on the whole path. After each failure, we can not have more than $N-1$ repairs on $[\tau_F < \tau_{\mathbf{1}}]$. We can not then have more than $t(N-1)$ repairs on the whole path, so, the total number of transitions can not be greater than $tN+1$. Thus

$$|\Delta'_t| \le |S|^{tN+1}.$$

The first inequality of the second part of the lemma is demonstrated in [8]. If ε is sufficiently small such that $\varepsilon\delta < 1$, as $t \le m$ (because $t = m - k$), we obtain the second inequality.

For the third part, we have

$$L(x_0, \cdots, x_n) = \prod_{k=0}^{n-1} \frac{\boldsymbol{P}(x_k, x_{k+1})}{\boldsymbol{P}'(x_k, x_{k+1})}.$$

By definition of $\mathcal{I}$, there exists $\nu(x,y)$ independent of ε such that

$$\nu(x,y)\boldsymbol{P}(x,y) \le \varepsilon\boldsymbol{P}'(x,y) \tag{2}$$

for all $(x,y) \in \Gamma$ $x \ne \mathbf{1}$, $y \prec x$ and all sufficiently small ε, and

$$\nu(x,y)\boldsymbol{P}(x,y) \le \boldsymbol{P}'(x,y) \tag{3}$$

for all $(x,y) \in \Gamma$, $y \succ x$ or $(\mathbf{1}, y) \in \Gamma$ and all sufficiently small ε. Let $\nu' = \min\{\nu(x,y) : (x,y) \in \Gamma,\ y \prec x\}$ and $\nu_* = \min(1, \nu')$. For all ε small enough,

$$L(x_0, \cdots, x_n) \le \frac{1}{\nu_*} \prod_{k=1}^{n-1} \frac{\varepsilon 1_{[(x_k, x_k+1) \in \Gamma,\ x_{k+1} \prec x_k]} + 1_{[(x_k, x_{k+1}) \in \Gamma,\ x_{k+1} \succ x_k]}}{\nu_*} \le \frac{\varepsilon^t}{\nu_*^{tN+1}}$$

because there are less that $t+1$ failures (one from state **1** for which the likelihood is bounded by $1/\nu^*$ by (3) and not more than t from other states for which the likelihood is bounded using (2) by ε/ν^*) and $t(N-1)$ repairs (likelihood is bounded using (3) by $1/\nu^*$). If we take $\eta = 1/\nu_*^N$, and $\kappa = 1/\nu_*$ we obtain the desired inequality. ∎

Proof of the Theorem:

Necessary condition: Suppose that there exist $k, m \in \mathbb{N}$ and $(x_0, \cdots, x_n) \in \Delta_{m,k}$ such that $k \geq 3m/2 - 3s/4 + 1$ and $m - k < r$. This means that

$$\Phi'\{(X_0, \cdots, X_{\tau_F}) = (x_0, \cdots, x_n)\} = O(\varepsilon^{3m/2-3s/4+1}).$$

Let

$$L(y_0, \cdots, y_n) = \frac{\Phi\{(X_0, \cdots, X_{\tau_F}) = (y_0, \cdots, y_n)\}}{\Phi'\{(X_0, \cdots, X_{\tau_F}) = (y_0, \cdots, y_n)\}}.$$

We have then

$$\begin{aligned}\rho &= \sum_{(y_0,\cdots,y_n)\in\Delta} |L(y_0, \cdots, y_n) - \gamma|^3 \, \Phi'\{(X_0, \cdots, X_{\tau_F}) = (y_0, \cdots, y_n)\} \\ &\geq |L(x_0, \cdots, x_n) - \gamma|^3 \, \Phi'\{(X_0, \cdots, X_{\tau_F}) = (x_0, \cdots, x_n)\} \\ &= \Theta(\varepsilon^{3(m-k)+k}) \\ &= \underline{O}(\varepsilon^{3s/2-2}).\end{aligned}$$

Thus $\rho/\sigma^3 = \underline{O}(\varepsilon^{-2}) \to +\infty$.

Sufficient condition: Suppose that

$$\Phi'\{(X_0, \cdots, X_{\tau_F}) = (x_0, \cdots, x_n)\} \ = \ \underline{O}(\varepsilon^{3m/2-3s/4})$$

for all $(x_0, \cdots, x_n) \in \Delta_{m,k}$, $m \geq r$ and $m - k < r$. We have to show that $\rho = O(\varepsilon^{3s/2})$.

$$\rho = \sum_t \sum_{(x_0,\cdots,x_n)\in\Delta'_t} |L(x_0, \cdots, x_n) - \gamma|^3 \Phi'\{(X_0, \cdots, X_{\tau_F}) = (x_0, \cdots, x_n)\}.$$

Let $(x_0, \cdots, x_n) \in \Delta'_t$ such that $t < r$ (i.e., $m - k < r$).

$$\begin{aligned}|L(x_0, \cdots, x_n) - \gamma|^3 \Phi\{(X_0, \cdots, X_{\tau_F}) = (x_0, \cdots, x_n)\} &= \frac{\Theta(\varepsilon^{3m})}{\Theta(\varepsilon^{3k})}\Theta(\varepsilon^k) \\ &= \frac{\Theta(\varepsilon^{3m})}{\underline{O}(\varepsilon^{2(3m/2-3s/4)})} \\ &= O(\varepsilon^{3s/2}).\end{aligned}$$

As $\sum_{t<r} |\Delta'_t| < +\infty$ and by the first part of Lemma 1,

$$\sum_{t<r} \sum_{(x_0,\cdots,x_n)\in\Delta'_t} |L(x_0,\cdots,x_n)-\gamma|^3 \Phi'\{(X_0,\cdots,X_{\tau_F})=(x_0,\cdots,x_n)\} = O(\varepsilon^{3s/2}).$$

Furthermore

$$\begin{aligned}
&\sum_{t=r}^{\infty} \sum_{(x_0,\cdots,x_n)\in\Delta'_t} |L(x_0,\cdots,x_n)-\gamma|^3 \Phi'\{(X_0,\cdots,X_{\tau_F})=(x_0,\cdots,x_n)\} \\
&\le \sum_{t=r}^{\infty} \sum_{(x_0,\cdots,x_n)\in\Delta'_t} (\gamma^3+3\gamma^2\kappa\eta^t\varepsilon^t+3\gamma\kappa^2\eta^{2t}\varepsilon^{2t}+\kappa^3\eta^{3t}\varepsilon^{3t})\alpha\delta^t\varepsilon^t \\
&\le \sum_{t=r}^{\infty} |S|^{tN}(\gamma^3+3\gamma^2\kappa\eta^t\varepsilon^t+3\gamma\kappa^2\eta^{2t}\varepsilon^{2t}+\kappa^3\eta^{3t}\varepsilon^{3t})\alpha\delta^t\varepsilon^t \\
&= \gamma^3\alpha\sum_{t=r}^{\infty}(\delta\varepsilon|S|^N)^t + 3\alpha\gamma^2\kappa\sum_{t=r}^{\infty}(\eta\delta\varepsilon^2|S|^N)^t + 3\alpha\gamma\kappa^2\sum_{t=r}^{\infty}(\eta^2\delta\varepsilon^3|S|^N)^t \\
&\quad +\alpha\kappa^3\sum_{t=r}^{\infty}(\eta^3\delta\varepsilon^4|S|^N)^t \\
&= \Theta(\varepsilon^{4r})+\Theta(\varepsilon^{4r})+\Theta(\varepsilon^{4r})+\Theta(\varepsilon^{4r}) \\
&= O(\varepsilon^{3s/2})
\end{aligned}$$

because $2r \ge s$. Thus the normal approximation is bounded. ■

4 Difference between bounded normal approximation and bounded relative error

For all measures $\Phi' \in \mathcal{I}$, by Theorem 2 of [8], we have a bounded relative error if and only if $\forall m$ such that $r \le m \le 2r-1$, and $\forall (x_0,\cdots x_n) \in \Delta_m$,

$$\Phi'\{(X_0,\cdots,X_{\tau_F})=(x_0,\cdots,x_n)\} = \underline{O}(\varepsilon^{2m-2r}).$$

In the following theorem we prove that the class of measures with bounded normal approximation is included in the class of measures having bounded relative error.

Theorem 5 *Consider a measure $\Phi' \in \mathcal{I}$. If we have bounded normal approximation, we have bounded relative error.*

Proof: Suppose that we have a bounded normal approximation. By definition of s there exists at least one $m \geq r$ and one path $(x_0, \cdots, x_n) \in \Delta_m$ such that

$$\frac{\Phi^2}{\Phi'}\{(X_0, \cdots, X_{\tau_F}) = (x_0, \cdots, x_n)\} = \Theta(\varepsilon^s),$$

thus

$$\Phi'\{(X_0, \cdots, X_{\tau_F}) = (x_0, \cdots, x_n)\} = \Theta(\varepsilon^{2m-s}).$$

As we have bounded normal approximation, $2m-s \leq 3m/2-3s/4 = 3/4(2m-s)$, then $2m - s \leq 0$. Then

$$\Phi'\{(X_0, \cdots, X_{\tau_F}) = (x_0, \cdots, x_n)\} = \Theta(1)$$

because we have always $2m \geq s$.

Let $(y_0, \cdots, y_n) \in \Delta_r$ and r' be such that $\Phi'\{(X_0, \cdots, X_{\tau_F}) = (y_0, \cdots, y_n)\} = \Theta(\varepsilon^{r'})$. By definition of r,

$$2r \geq 2r - r' \geq s = 2m \geq 2r$$

with the previous m, then

$$s = 2r.$$

Suppose that we have not bounded relative error. There exists $(z_0, \cdots, z_n) \in \Delta$, with $\Phi\{(X_0, \cdots, X_{\tau_F}) = (z_0, \cdots, z_n)\} = \Theta(\varepsilon^l)$ and $\Phi'\{(X_0, \cdots, X_{\tau_F}) = (z_0, \cdots, z_n)\} = \Theta(\varepsilon^{l'})$, such that $l' > 2l - 2r$, which means that $l' > 2l - s$. But by definition of s, $s \leq 2l - l'$. We have then proved the theorem. ∎

We can find a system example with bounded relative error, but without bounded normal approximation. Suppose that $C = 2$, $n_1 = n_2 = 2$ and that the system is operational if at least two components are operational. Let the transitions of the DTMC of this system be represented by the transitions on Figure 1, where the failed states are filled.

If we use as importance sampling scheme Bias1, with the new probabilities described in Figure 2, then we have a bounded relative error, but not a bounded normal approximation. This means that, even if we have bounded relative error, the

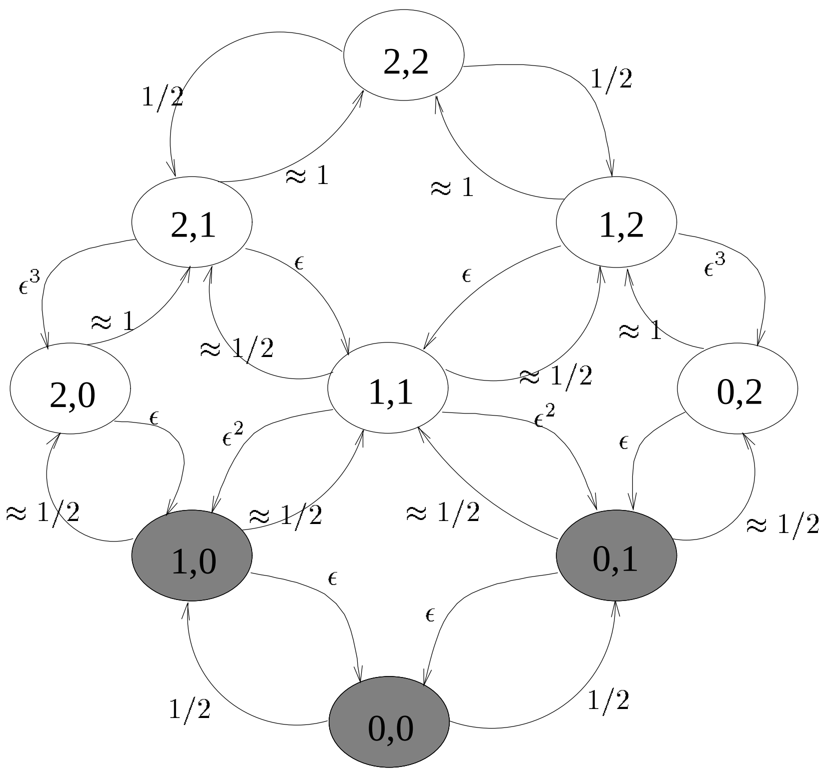

Figure 1: Transitions of system I

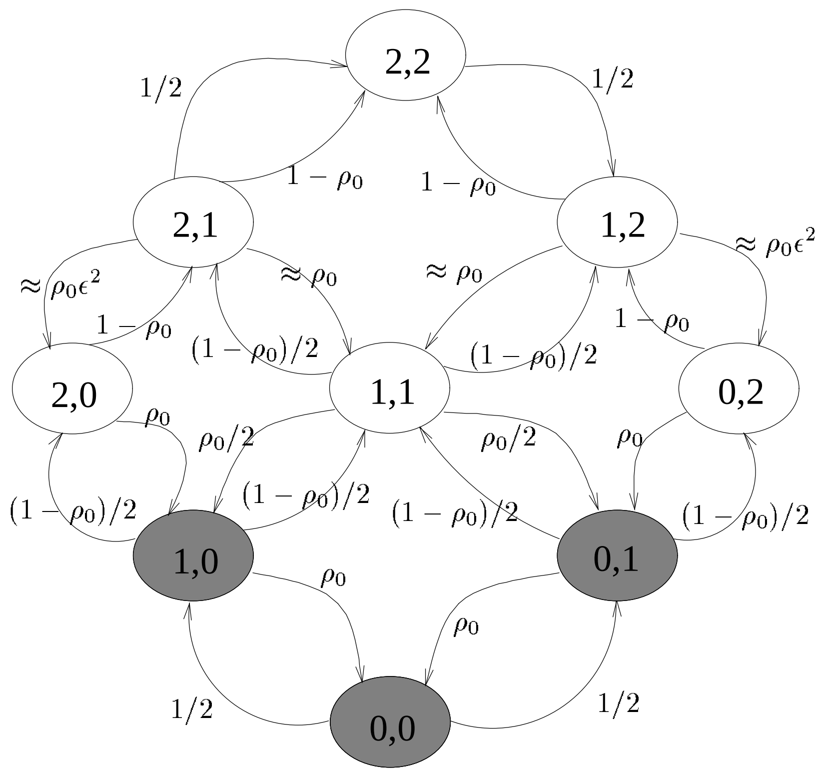

Figure 2: Bias1 importance sampling transitions of system I

confidence interval is not valid for high reliability values. For this example,

$$E_{\Phi}(1_{[\tau_F<\tau_{\mathbf{1}}]}) = 2\varepsilon^3 + o(\varepsilon^3),$$

$$\sigma^2 = \frac{5}{\rho_0^2}\varepsilon^6 + o(\varepsilon^6)$$

and

$$\rho = \frac{1}{\rho_0^4}\varepsilon^8 + o(\varepsilon^8).$$

For this system, since there is no propagation fault, Bias1 and distance importance sampling measures are equivalent, so, this is also an example of the distance importance sampling technique giving bounded relative error but not bounded normal approximation. We can also exhibit systems (with at least three types of components) with the same property for the Bias2 technique.

Theorem 6 *Balanced failure biasing has the property of bounded normal approximation. Simple failure biasing, Bias2 failure biasing and distance biasing, which in general do not have bounded relative error, do not verify the bounded normal approximation. Nevertheless, for balanced systems (i.e. systems for which failure transitions have probabilities of the same order), all the methods give bounded normal approximation.*

Proof: The proof that balanced failure biasing for every system give bounded normal approximation results directly from the necessary and sufficient condition given in Theorem 4 and the fact that $\Phi'\{(X_0, \cdots, X_{\tau_F}) = (x_0, \cdots, x_n)\} = \Theta(1)$ (i.e. $k = 0$) for all paths to failure $(x_0, \cdots, x_n) \in \Delta$, because each probability in matrix $\boldsymbol{P'}$ is in $\Theta(1)$. The same arguments work in the case of balanced systems. Counter-examples for Bias1 and the distance-based technique are given above. A counter-example for Bias2 can also be built in a similar manner. ∎

5 Conclusion

The objective of this paper is to define the concept of bounded normal approximation and to underline its importance in the context of the evaluation of dependability measures using Markov models. Then we give a necessary and sufficient condition to obtain a bounded normal approximation in highly reliable Markovian systems. Up to now, literature has focused on bounded relative error. A good importance sampling measure should verify both properties. But if we have bounded normal approximation, we have bounded relative error. Thus our necessary and sufficient condition is more attractive. We have also proven that for the usual importance sampling schemes (Bias1, Bias2 and the distance-based technique), it is possible to have bounded relative error without bounded normal approximation. Nevertheless, on balanced systems, we always have bounded normal approximation. The advantages of balanced failure biasing with respect to the other schemes are thus reinforced by the results of this paper.

References

[1] J. A. Carrasco. Failure distance based on simulation of repairable fault tolerant systems. *Proceedings of the 5th International Conference on Modelling Techniques and Tools for Computer Performance Evaluation*, pages 351–365, 1992.

[2] W. Feller. *An Introduction to Probability Theory and its Applications*, volume II. John Wiley and Sons, 1966. second edition.

[3] A. Goyal, L. Lavenberg, and K. Trivedi. Probabilistic Modeling of Computer System Availability. *Annals of Operations Research*, 8:285–306, 1987.

[4] A. Goyal, P. Shahabuddin, P. Heidelberger, V. F. Nicola, and P. W. Glynn. A Unified Framework for Simulating Markovian Models of Highly Dependable Systems. *IEEE Transactions on Computers*, 41(1):36–51, January 1992.

[5] J. M. Hammersley and D. C. Handscomb. *Monte Carlo Methods*. Methuen, London, 1964.

[6] E. E. Lewis and F. Bohm. Monte Carlo Simulation of Markov Unreliability Models. *Nuclear Engineering and Design*, 77:49–62, 1984.

[7] R.R. Muntz, E. de Souza e Silva, and A. Goyal. Bounding Availability of Repairable Computer Systems. *IEEE Transactions on Computers*, 38(12):1714–1723, 1989.

[8] M. K. Nakayama. General Conditions for Bounded Relative Error in Simulations of Highly Reliable Markovian Systems. Technical report, IBM Research Division, June 1993.

[9] P. Shahabuddin. Importance Sampling for the Simulation of Highly Reliable Markovian Systems. *Management Science*, 40(3):333–352, March 1994.

Unit´e de recherche INRIA Lorraine, Technopˆole de Nancy-Brabois, Campus scientifi que,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unit´e de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unit´e de recherche INRIA Rhˆone-Alpes, 655, avenue de l'Europe, 38330 MONTBONNOT ST MARTIN
Unit´e de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unit´e de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
ISSN 0249-6399